Real time face detection and recognition

Ashik Mahmud Shawon

ID: 2014-2-60-110

A thesis submitted in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science and Engineering

Department of Computer Science and Engineering
East West University
Dhaka-1212, Bangladesh

August, 2018

Declaration

I, hereby, declare that the work presented in this thesis is the outcome of the research performed by me under the supervision of Dr. Taseed Jabid, Assistant Professor, Department of Computer Science and engineering, East West University. I also declare that no part of this thesis has been or is being submitted elsewhere for the award of any degree or diploma.

Countersigned	Signature
. . . . .	. . . . .
(**Dr. Taskeed Jabid**)	(**Ashik Mahmud Shawon**)
Supervisor	(ID: 2014-2-60-110)

Letter of Acceptance

This thesis report entitled "**Real time face detection and recognition**" submitted by "Ashik Mahmud Shawon" to the Department of Computer Science and Engineering, East West University is accepted by the department in partial fulfillment of requirements for the Award of the Degree of Bachelor of Science and Engineering on September, 2018

Supervisor	Chairperson
…….	……
(Dr. Taskeed Jabid)	(Dr. Ahmed Wasif Reza)
Assistant Professor	Associate Professor and chairperson
Department of CSE	Department of CSE
East West University	East West University

Abstract

Performing face detection in real time is a challenge as accuracy and efficiency both are highly considerable here. In real time face detection system has to face many difficulties as it is more difficult to extract useful features from a moving object. There are different types of method to detect and recognize human faces. In this thesis, I am trying to extract feature from face image at real time and proposed a model that can detect human faces and recognize them. First I use Haarcasecade classifier to classify and detect human faces then I use LBPH algorithm to extract feature from dynamic image and compare them with training data to recognize them.

Acknowledgments

As it is true for everyone, I have also arrived at this point of achieving a goal in my life through various interactions with and help from other people. However, written words are often elusive and harbor diverse interpretations even in one's mother language. Therefore, I would not like to make efforts to find best words to express my thankfulness other than simply listing those people who have contributed to this thesis itself in an essential way. This work was carried out in the Department of Computer Science and Engineering at East West University, Bangladesh.

First of all, I would like to express my deepest gratitude to Allah for blessings on us. Next, my special thanks go to my supervisor, "Dr. Taskeed Jabid", who gave me this opportunity, initiated me into the field of "Face Detection and Recognition Using Haar feature and LBPH", and without whom this work would not have been possible. His encouragements, visionaries and thoughtful comments and suggestions, unforgettable support at every stage of my BSc study were simply appreciating and essential. His ability to muddle me enough to finally answer my own question correctly is something valuable what I have learned and I would try to emulate, if ever I get the opportunity.

Last but not least, I would like to thank my family members for supporting us. There are numerous other people too who have shown me their constant support and friendship in various ways, directly or indirectly related to my academic life. I will remember them in my heart and hope to find a more appropriate place to acknowledge them in the future.

Ashik Mahmud Shawon

April, 2018

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Face Detection and Recognition

Every human faces got different structure. Analyzing the facial features faces can be detected from image. The main task is analyzing the pixels and find out some unique features that represent faces. After detection recognition is another important part. Recognizing the faces this is the most challenging task to do in image processing. This actually measuring a unique pattern from training faces and compares it with other test images to recognize the face. The most accurate feature or pattern will give me most accurate result. [1]

1.1.1 Why face recognition is important?

Face recognition is a biometric identification. If system can recognize every human that can be used for security purposes [2]. As I see now-a-days security camera is available almost everywhere. It is just watching what is happening but it can't act on certain situation. If faces that caught on camera can be recognizable that can help in many ways such that employee attendance, class participation, or any security purposes. If I train the systems with criminal images collected from local police it's add extra security alert when these faces will be detected. This will be work as prevention for the security agency that they would know who could be make occurrence and security agencies could be alert for any situation that time.

1.2 Different approaches of face recognition

Face recognition is a challenge for now-a-days. So there are some methods that is used frequently for face recognition or facial expression recognition in image processing. There are 3 methods[3] that are mostly used worldwide. Those are:

Input Image Normalization[3]
Geometric Approach[3]
Neural Networks for Access Control[3]

1.2.1 Geometric Method

Geometric method is a useful approach for facial expression recognition. In this approach the first step is to obtain a training set. After that it calculates the histogram of the images to detect the limit of the skin region. Then some pre-processing is needed like erosion, dilation, and filtering. After applying this skin region of the face is found which describe the geometric shape of the face. After that I can find the edges of the head, to enter them to process of finding angles [4]. This method is quite effective but it needs a lot of pre-processing.

After applying all those pre-work, only then I can apply this method. Also if I make any mistake in pre-process then the recognition system will fall apart if the Geometric algorithm is right though. For avoiding all those pre-processing the other method is preferable to use, that is Global method. One of them is Local Binary Pattern or LBP.

1.3 Local Binary Pattern

Local Binary Pattern (LBP) is an efficient texture operator which labels the pixel of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. [5]

1.3.1 Process of LBP

At first, I have to divide the whole image into cells. [6]

Then for each pixel in a cell, I have to compare the pixel to each of its 8 neigh-bors. I have to follow the pixels along a circle for example clockwise or counter-clockwise. [6]

”0” is written when the center pixel’s value is greater than the neighbor’s value, otherwise ”1” is written. This gives an 8-digit binary number. [6]

We have to compute the histogram, over the cell, of the frequency of each” number” occurring. This histogram can be seen as a 256-dimensional feature vector. [6]

Then I have to concatenate histograms of all cells. This gives a feature vector for the entire window. [6]

Mathematically,

$$LBP_{R;P} = \sum_{P=1}^{P\sum 1} s(g_P \quad g_C).2^P \ldots\ldots\ldots\text{(i)}$$

Where g_P is neighborhood pixel in each block and it is threshold by its center pixel value(g_C).P is a sampling point and R is radius.

Binary threshold function is,

$$S(x) = \begin{matrix} 0, x & 0 \\ 1, x & 0 \end{matrix} \ldots\ldots\ldots\ldots\ldots\ldots\text{(ii)}$$

In local binary pattern or LBP the size of features increases exponentially with the number of neighbors. This excessive number of features can be very tough to handle. It increases the complexity of the system. The structural information in here is limited be-cause only pixel difference is used, magnitude information is ignored [7]. So for ignoring

a huge number of features and a large array in code I moved into another algorithm where a lot of features is not needed. The algorithm is called Principal Component Analysis or PCA in short.

1.4 HaarCascade Face Detection

.

Haar Features that is based on cascade classifiers is an effective object detection method. It is first introduced by Paul voila and Michael Jones in 2001 [8]. This approach is basically a machine learning approach. Cacsecade Function takes trainings from a lots of positive and negative images. This training experience is used for taking decisions at test image. After training I need to extract feature from it. For this following images below are used. Every feature is just a single value that is the difference of dark pixels and bright pixels [9].

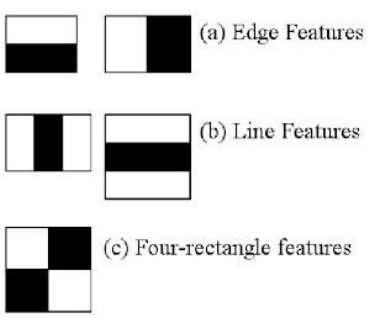

Figure 1.1: Windows uses to extract haar features

But there are many number of features is available in an image that is needed to extracted. If window size is 24x24 that would be needed 160000 features to be extracted. For each feature calculation I need to subtract the sum of white pixel and black pixel. To solve this problem, they introduced integral image [8]. It reduced calculations so this is the ultimate solution for image size. It also makes calculations super faster. but all of the calculated feature is not important even most of them are irrelevant. Let's consider the image below. Top rows of the images

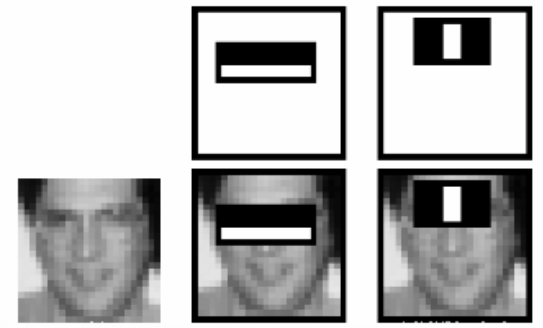

Figure1.2: extracting features from image with windows

shows good features. The first feature selected seems to focus on the property that region of eye is often darker than region of forehead, region of cheeks and region of nose. The second feature selected relies on the property of eyes are darker than the bridge of nose. But when the same window applies on cheeks or many places is irrelevant. So there is a problem to select best features from a lot of features and it is achieved by Adaboost. After applying Adaboost features reduced at 6000. But this is still a big amount of data and this is time consuming. For this authors introduced the concept of casecade of classifiers for time efficiency.

1.4.1 Adaboost

For this, I apply each and every feature on all the training images. For each feature, it finds the best threshold which will classify the faces to positive and negative. Obviously, there will be errors or misclassifications. I select the features with minimum error rate, which means they are the features that most accurately classify the face and non-face images. (The process is not as simple as this. Each image is given an equal weight in the beginning. After each

classification, weights of misclassified images are increased [10]. Then the same process is done. New error rates are calculated. Also new weights. The process is continued until the required accuracy or error rate is achieved or the required number of features are found The final classifier is a weighted sum of these weak classifiers. It is called weak because it alone can't classify the image, but together with others forms a strong classifier. The paper says even 200 features provide detection with 95% accuracy.

1.4.2 Integral Image

Rectangle features can be computed very rapidly using an intermediate representation for the image which I call the integral image.3 The integral image at location (x,y) contains the sum of the pixels left and above[8]

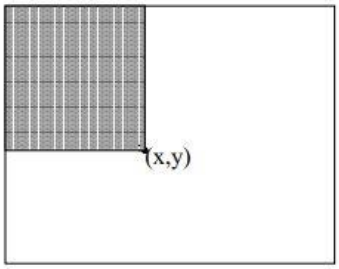

Figure 1.3: The value of the integral image at point (x,y)

the left of (x,y) inclusive:

$$ii(x,y) = \sum_{x' \le x, y' \le y} i(x', y'),$$ ……(iii)

This actually helps to optimize calculation of haar features with viola-jones algorithm [8].

1.4.3 Cascade of Classifiers

Authors divides all features after minimizing by adabost classifier. Instead of put every window on every features they use different window on different cascades. If the window fails to extract useful feature for first calculation, then it will be ignored. Program doesn't consider rest other features for this cascade. And if it passes apply the second stage of features and continues the process [9].

1.5 Organization of the Book

We have structured my rest of the Thesis works as following: In Chapter 2, my working process is discussed broadly in Chapter 3; Chapter 4 introduces the experimental result of my working process. Chapter 5 represents the conclusion of my work and proper references of my thesis works.

Chapter 2

Related Works

2.1 Introduction

At technological level, image processing is a new era of development. This field is so promising now-a-days. There are a lots of opportunities here. One field of image processing is face recognition. It has a huge opportunity in various field. One of them is facial expression recognition. There's a lot of scope in this field like improving human-computer interaction, mood detection using facial expression, controlling the uses of different kind of apps on the basis of anyone's mood or controlling the environment facilities in the basis of someone's mood. So this can make my life easy.

2.2 Related Works

various authors tried to represent their thoughts and experiments about facial expression recognition which are as follows:

2.2.1 Local gradient hexa pattern: a descriptor for face recognition and retrieval.

Soumwndu Chakraborty, Satish Kumar Sing, Pavan chakraborty describe this method in their paper headings with “Local gradient hexa pattern: a descriptor for face recognition and retrieval” [11]. It identifies variations in the original pixels of a facial image in higher order derivative space. It has it’s won encoding function and structure of micropattern. First order derivative for first order directional derivatives of image size M x N for distance D(distance from center point) and angle alpha(0,45,90,135). This process will be continue further derivative where D will be calculated from micropattern. This process will extract features from image to detect the face. But this process is not lightweight to use real time. This can detect face on static image.

2.2.2 A face recognition process based on LBP feature for CNN.

In this paper [12] authors used LBP to extract features form image but they neglected position information of LBP. LBP features is Most Basic unit of texture detection of image. It uses the statistical method to detect faces and misses other important features. But face recognition based on CNN take image as input and CNN learns from pixel levels. In this learning process of CNN The feature extraction start from lowest level feature. Accuracy of this model is good. It is 95.33%. [12] But applying CNN in real time is quite difficult.

2.2.3 A new GLBSIF Descriptor For Face Recognition.

In this paper [13] the authors proposed a method of three processes combine together. They called it Hybrid Descriptor. For this they use a normalized image instead of original image. After normalizing the image, they use Gabor filters. After filtering the image LBP applied to extract the features. Then they use Binarized Statistical image features(BSIF). BSIF is a new descriptor. A binary code for each pixel was computed using linearly projecting of the image local patches onto a subspace whose main vectors are learned from natural images by independent component analysis. Applying BSIF on image results more representing statistical information. After applying GLBSIF they use PCA and then KMN-SRC classification to recognize the face. Accuracy is good here and it is 97.81%.

2.2.4 Face Recognition using Eigenfaces.

In here [14], Acquire a database of face images, calculate the eigenfaces a determine the face space with all them. It will be necessary for further recognitions. When a new image is found, calculate its set of weights. Determine if the image is a face; to do so, I have to see of it is close enough to the face space. Finally, it will be determined if the image corresponds to a known face of the database or not. Since the objective of this project is not the implementation of the algorithms, but the description and comparison of them, the results will be reported from the experiments performed by the authors of the articles. Both methods were tested using variations of face orientation, illumination and size.

2.2.5 Face recognition using eigen-faces, fisher-faces and neural networks.

In this paper [15], In this paper, a new face recognition method based on PCA (principal component analysis), LDA (linear discriminant analysis) and neural networks is proposed. This method consists of four steps: i) preprocessing, ii) dimension reduction using PCA, iii) feature extraction using LDA and iv) classification using neural network. Combination of PCA and LDA is used for improving the capability of LDA when a few samples of images are available and neural network classifier is used to reduce number misclassification caused by not-linearly separable classes. The proposed method was tested on Yale face database. Experimental results on this database demonstrated the effectiveness of the proposed method for face recognition with less misclassification in comparison with previous methods.

Chapter 3

Methodology

3.1 Classifier Training

To recognize a face first system needs to train the classifier with specific image of face. During training system actually extract feature from it or make a histogram from it. System takes average of all features. To take train data it is considerable to take only face images. I take face images from various angels and expressions from the data make a train file that will help to recognize input. That increases recognition accuracy.

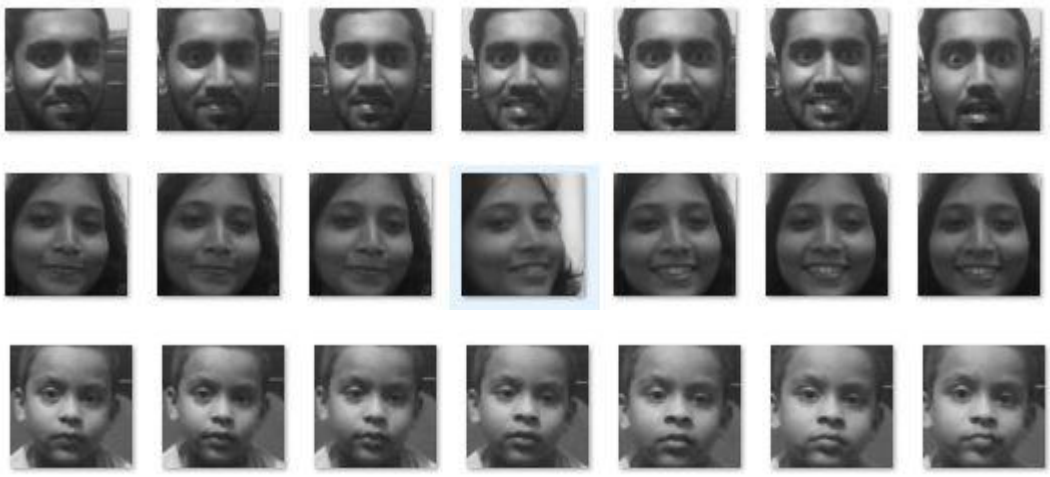

Image 3.1: Training data set

3.2 Taking inputs

I'm here working on image processing. I exactly am doing face recognition. To taking input for real time I use a camera that can record video images. System takes every frame of video as a image.

3.2 Detecting Faces With Haar Casecade Classifier

After taking image the very first thing is need to do is detect the face. I use Haar cascade classifier to detect the faces from image.

3.3 Applying LBPH

After detecting the faces, the main task to do is applying LBPH algorithm to extract features. Divide the image into cells (suppose 256*256). Then I compare each cells pixel value with its 8 neighbor cells either clockwise or anti-clockwise. If the center pixel value is greater than its neighbors value, then I put 0 otherwise 1. By doing so I get

an 8-digit binary number and then I convert it into decimal and put it in the center cell. After doing this I get a 254*254 image. Then I calculate the histogram over the cell which is the frequency of each number occurring in the image. Then I put all the histogram value in a file. To recognize a face first system needs to train the classifier with specific image of face. During training system actually extract feature from it or make a histogram from it. System takes average of all features. To take training data it is considerable to take only face images. At this point, now I am trying to work with features. Because I want to improve the recognition system. For this I divide my images into 6 different blocks and then I apply the LBP algorithm on the blocks of images and get LBP histogram for the blocks on images. After that I put the histogram values along with images number in files again.

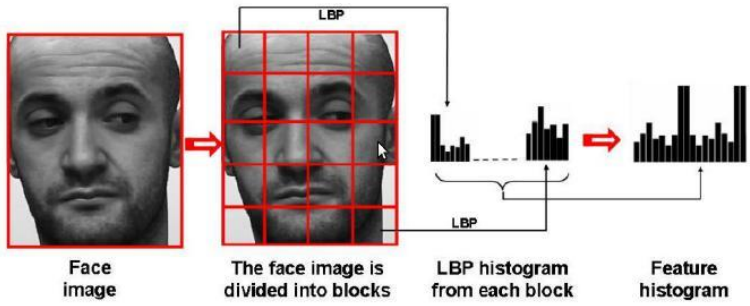

Figure 3.2: Local Binary pattern

After creating histogram system will compare the histogram with the histogram that is created during the train period. Based on this matching result system could predict the person if it knows the face or not.

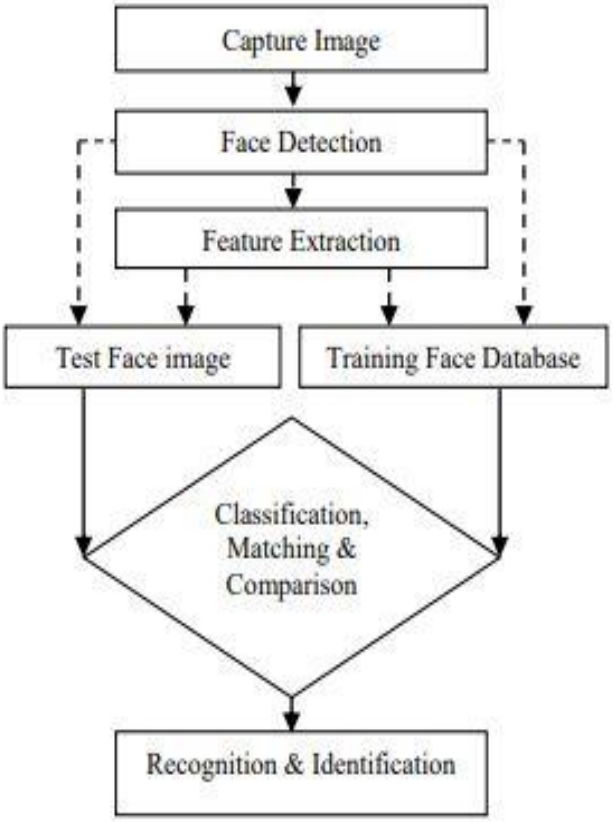

Figure 3.3: flow chart of face recognition

Chapter 4

Results

4.1 Results

This process results with the face that is known or unknown if the process recognizes the faces it will replies with name.

Image:4.1 output window

4.2 Accuracy

In this real case scenario accuracy is not that much high as algorithm says for static image situation. In real case there can be may situations. Image can be come out from many different angles. Difference of light is existed there. Distance from camera is also a fact for this difference in histogram that cost low accuracy of system.

4.2.1 Accuracy for Face Detection

Calculating Accuracy for detecting face is not same for all angles. Accuracy for frontal face and turned face is not same. Feature extracted from frontal face is more informative than turned faces. So It is easy to detect a frontal face with Haar features. Turned face is also detectable but if it is full turned then face will be un detectable.

4.2.2 Accuracy for Face Recognition

To recognizing a face system needs more information that to detect it. So frontal face is almost always recognizable. But for turned face proper comparable features lacks. So face is unrecognizable when it positioning fully turned.

4.2.3 Factors That Hampers Accuracy While Performing In Real Time

The factors that hampers accuracy for performing in real time:

- Algorithmic Limitations.
- Picture Quality.
- Amount Of Light.
- Position of face.

Chapter 5

Future work & Conclusion

5.1 Future Work

Today, one of the fields that uses facial recognition the most is security. Facial recognition is a very effective tool that can help law enforcers recognize criminals and software companies are leveraging the technology to help users access their technology. This technology can be further developed to be used in other avenues such as ATMs, accessing confidential files, or other sensitive materials. This can make other security measures such as passwords and keys obsolete.

Another way that innovators are looking to implement facial recognition is within subways and other transportation outlets. They are looking to leverage this technology to use faces as credit cards to pay for your transportation fee. Instead of having to go to a booth to buy a ticket for a fare, the face recognition would take your face, run it through a system, and charge the account that you've previously created. This could potentially streamline the process and optimize the flow of traffic drastically. The future is here.

5.2 Conclusion

In this thesis study, I applied Haar Cascade and LBPH algorithm for facial detection and recognition. First I used Haar feature to detect face. After that I applied LBP feature and make histogram with it of an image. Then I compared histogram with training data. Then based on comparison result I can recognize the face. Accuracy is not so good because of this is a real case scenario. Collecting training data for system is a challenge because face can be appearing on camera in various angles and different light condition that differs in extracted features. If I can detect faces and recognize them correctly, then I will be able to develop much automatic system which will develop their program based on authentic face. The whole world will be smarter than ever and my life will be a lot easier. So my study can be helpful for the researcher who are continuously working for the development of face recognition and working for a better world. Rectangle features can be computed very rapidly using an intermediate representation for the image

Bibliography

[1]. Zahid Riaz, Christoph Mayer, Michael Beetz, and Bernd Radig,**" *Model Based Analysis of Face Images for Facial Feature Extraction*".** Available: https://pdfs.semanticscholar.org/7b8e/2b1f759424a0e909095c7aae3727377ffbb4.pdf

[2]. Mrs. Sunita Roy and Mr. Susanta Podder, **"*Face detection and its applications*"**, IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 1, Issue 2, April-May, 2013, ISSN: 2320 – 8791

[3]. V.V. Starovoitov , D.I Samal , D.V. Briliuk, ***"THREE APPROACHES FOR FACE RECOGNITION",*** The 6-th International Conference on Pattern Recognition and Image Analysis October 21-26, 2002, Velikiy Novgorod, Russia, pp. 707-711

[4]. Ayman EL-SAYED and Osama Rezq, "***Geometrical Approach for Face Detection and Recognition***".

[5]. M. Pieticaginen, ***"Local binary Pattern",*** FI-90014, University of Oulu, Finland**,** 2018.

[6]. "HTTPS", en.wikipedia.org, 2018. [online]. Available: https://en.wikipedia.org/wiki/HTTPS. [Accessed: 30-August-2018]

[7]. Biomia.org, 2018. [Online]. Available: http://biomisa.org/uploads/2016/10/Lect-15.pdf. [Accessed: 30-August-2018]

[8]. Paul Viola and Michael Jones, ***"Robust Real-time Object Detection".*** Second international workshop on statistical and computational theories of vision – modeling, learning, computing, and sampling vancouver, canada, july 13, 2001.

[9]. Opencv.org, "Face Detection using Haar Cascades" Available: https://docs.opencv.org/3.4/d7/d8b/tutorial_py_face_detection.html

[10]. Ruihu Wang, " AdaBoost for Feature Selection, Classification and Its Relation with SVM, A Review". 2012 International Conference on Solid State Devices and Materials Science. Available: https://www.sciencedirect.com/science/article/pii/S1875389212005767

[11]. Soumwndu Chakraborty, Satish Kumar Sing and Pavan Chakraborty, **"**Local gradient hexa pattern: a descriptor for face recognition and retrieval**",** ieee transactions on circuits and systems for video technology, vol. 28, no. 1, january 2018

[12]. Hongshuai Zhang, Zhiyi Qu, Liping Yuan and Gang Li, "A Face Recognition Method Based on LBP Feature for CNN", 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), March 2017.

[13]. Bilel Ameur, Mebarka Belahcene, Sabeur Masmoudi, Amira Guidara Derbel and Ahmed Ben Hamida, ***"A new GLBSIF Descriptor For Face Recognition"*** , 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), May 2017.

[14]. M.A. Turk and A.P. Pentland, ***"Face recognition using eigenfaces",*** Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 3-6 June 1991.

[15]. Hossein Sahoolizadeh and Youness Aliyari Ghassabeh, ***"Face recognition using eigen-faces, fisher-faces and neural networks"***, 2008 7th IEEE International Conference on Cybernetic Intelligent Systems, 9-10 Sept. 2008.